

Hero Elementary Playlist Pilot Study in After-school Programs

Linlin Li, Ph.D.

Ben Mahrer

Gary Weiser, Ph.D.

Ari Orenstein

Eunice Chow

Sara Atienza

August 2019

© 2019 WestEd. All rights reserved.

Suggested citation: Li, L., Maher, B., Weiser, G., Orenstein, A., Chow, E., & Atienza, S. (2019). *Hero Elementary playlist pilot study in after-school programs*. San Francisco, CA: WestEd.

WestEd is a nonpartisan, nonprofit research, development, and service agency that works with education and other communities throughout the United States and abroad to promote excellence, achieve equity, and improve learning for children, youth, and adults. WestEd has more than a dozen offices nationwide, from Massachusetts, Vermont, Georgia, and Washington, DC, to Arizona and California, with headquarters in San Francisco.



Contents

Study Overview	1
Description of Playlists	1
Study Design and Methodology	4
Treatment and Comparison Conditions	4
Instruments	5
Outcome Measures	5
Implementation Measures	6
Study Sample Description	7
Data Analysis Methods	8
Playlist Implementation in After-School Programs	9
Playlist Completion	9
Student Behavior and Affect	10
Classroom Integration	11
Connections	13
Classroom Management	13
Age Appropriateness of Content	14
Technical issues	16
Impact Results	19
Student Science Performance	19
Student Science Attitudes	19
Conclusion	21
References	21
Appendices	23

Study Overview

Schools that serve a majority of economically disadvantaged students often struggle with limited budgets to find the resources and educators for their science classrooms. Across the United States, issues of access and equity in Science, Technology, Engineering, and Math (STEM) education – tied intrinsically to fiscal resources – have led to the emergence of after-school programs rooted in youth development. As an informal learning setting, after-school programs are increasingly viewed as both complementary and supplementary to school learning (Pierce, Bolt, & Vandell, 2010). After-school programs have potential to affect students' in-school science achievement and interest in pursuing science careers. Students participating in highquality, STEM-related after-school programs have shown increased academic outcomes on standardized tests and a greater likelihood of pursuing a STEM-related career path during postsecondary studies than their peers who did not participate in such programs (Dabney et al., 2012; Vandell, Reisner, & Pierce, 2007). Further, educators in after-school programs are not constrained by the same state or local requirements as their regular school-day colleagues, affording after-school educators the opportunity to provide students with hands-on learning experiences or long-term projects that might not be possible during the regular school day (Peterson & Fix, 2007). In any case, effective after-school programs can have many benefits, particularly to students who otherwise may not have access to engaging and active learning opportunities (Basu & Barton, 2007; Dabney et al., 2012; Ozel et al., 2013).

Funded by the U.S. Department of Education, the Twin Cities Public Television (TPT) Ready to Learn project, *Hero Elementary*, has an emphasis on reaching Latino communities and supporting the needs of children with disabilities. The project embeds the expectations of kindergarten to 2nd grade science standards into a series of activities, including interactive games, educational apps, non-fiction e-books, hands-on activities, and a digital science notebook. The activities are organized into different playlists for educators and students to use in after-school programs, and each playlist centers on a meaningful conceptual theme in K-2 science learning.

In the spring of 2019, WestEd conducted a pilot study using five playlists to understand the feasibility of implementing the playlists in after-school programs and to discuss the potential impact of the playlists on student science learning. The following research questions guided this study:

- RQ1. How are the playlists implemented in after-school programs?
- RQ2. What is the potential impact of playlists on student science knowledge and skills?

Description of Playlists

The five playlists included in the spring pilot study were *Pushes and Pull; Make it Fast, Make it Slow; Changing Motion; Classifying Matter;* and *Solid or Liquid.* This section provides an

overview of the learning goals for each playlist, as well as a brief description of the activities therein.

Pushes and Pulls

In *Pushes and Pulls* (Table 1), students learn to distinguish a push from a pull and explore the ways that pushes and pulls can move an object. They discover how pushes and pulls can start an object moving when it is at rest and can bring a moving object to a halt. Students investigate little, medium, and big pushes and pulls.

Table 1. Pushes and Pulls activities

ACTIVITY	TYPE	LEARNING GOALS
Start, Stop, Go	Hands-On Activity	Explore how pushes and pulls make objects start, stop, and go at
Start, Stop, Go	Trands-On Activity	different speeds
Pushes and Pulls	Digital E-book	Introduce students to how pushes and pulls can make objects
Books!	Digital E-000k	change directions
How I Move Things	Digital Notebook	Students reflect on their experiences moving objects with big and
Tiow I wove Timigs	Digital Notcook	little pushes
With a Little Push!	Digital Episode	Observe how the Sparks Crew uses pushes and pulls to solve
With a Little I ush:	Digital Episode	problems
Push Power!	Digital Notebook	Describe the strength of a push required to move a heavy object
One More Push	Hands-On Activity	Investigate how objects of different weights move at different
One whole I ush	Tranus-On Activity	speeds when pushed
Push! Pull!! Puzzles!!!	Digital Game	Manipulate different kinds of objects using pushes and pulls

Make it Fast, Make it Slow

Students explore pushes and pulls in greater depth in *Make it Fast, Make it Slow* (Table 2). They explore the differences between pushes and pulls of varying strengths and they think about objects' speed.

Table 2. Make if Fast, Make it Slow activities

ACTIVITY	TYPE	LEARNING GOALS			
Make it Fast, Make it	Digital E-book	Explore basic concepts related to motion, including speed, force,			
Slow Books!	Digital E-000k	and inertia			
"X" Marks the Spot!	Hands-On Activity	Use pushes and pulls of varying strength to move objects in			
74 Warks the Spot.	Traines On 7 territy	different directions			
Move it Which Way	Digital Notebook	Record observations from the "X" Marks the Spot! activity			
A Soapy Situation	Digital Episode	Learn along with the Sparks Crew as they explore how the			
A Soapy Situation	Digital Episode	strength of a push can affect an object's speed			
Push! Pull!! Puzzles!!!	Digital Game	Use pushes and pulls of different strengths to manipulate objects			
Fast or Slow	Digital Notebook	Use predictions to explore how objects move with small pushes			
		Compare how pushes of different strengths affect the speed of an			
Lidtop Slide	Hands-On Activity	object; discover how pushes and pulls can move objects in			
		different directions			

Changing Motion

In *Changing Motion* (Table 3), students continue their exploration of how to manipulate and move objects. They discover the different ways that an object's speed and direction can change when it collides with another object. They also use cause and effect thinking to explore how to move objects toward a target.

Table 3. Changing Motion activities

ACTIVITY	ТҮРЕ	LEARNING GOALS
A Soapy Situation	Digital Episode	Learn along with the Sparks Crew as they explore how the
A Soapy Situation	Digital Episode	strength of a push can affect an object's speed
Push, Slide, Bump Pt. I	Hands-On Activity	Discover how moving objects interact when they bounce off one
Tush, Shue, Dump Tt. I	Trailus-On Activity	another
Changing Speed	Digital Notebook	Reflect on the Push, Slide, Bump activity; record observations
Duch Clide Duma Dt II	Hands-On Activity	Continue exploring how objects collide. Figure out how to control
Push, Slide, Bump Pt. II	Hands-On Activity	collisions with pushes and pulls of different strengths
Changing Motion	Digital E-book	Begin learning about forces and how objects move over different
Books!	Digital E-000k	surfaces (friction)
Changing Direction	Digital Notebook	Record the most effective strategy for moving objects through the
Changing Direction	Digital Notebook	Push, Slide, Bump obstacle course
Push, Slide, Bump Pt.	Hands-On Activity	Continue exploring how objects move and collide by building new
III	Tranus-On Activity	obstacle courses

Classifying Matter

In *Classifying Matter* (Table 4), students explore the different kinds of materials out of which objects are made. They make observations about the properties of objects and compare objects' similarities and differences. Students practice sorting materials into groups according to their different physical properties.

Table 4. Classifying Matter activities

ACTIVITY	TYPE	LEARNING GOALS
A Sorting Challenge	Hands-On Activity	Practice sorting objects by their different properties
The Right Stuff	Digital Episode	Observe how the Sparks Crew sorts materials according to their properties
Spot Search!	Hands-On Activity	Identify and observe properties of objects
My Favorite	Digital Notebook	Practice describing properties of matter
Scavenger Hunt	Hands-On Activity	Explore surroundings to find objects with certain properties
Sort and Group	Digital Notebook	Practice sorting materials and objects by their different properties
Classifying Matter Books!	Hands-On Activity	Learn the basic states of matter and become excited about hands- on scientific exploration

Solid or Liquid

Students learn about states of matter in the *Solid or Liquid* playlist (Table 5). They observe and describe solids and liquids and learn how to classify materials by state. Students also explore how to heat and cool matter to change its state.

Table 5. Solid or Liquid activities

ACTIVITY	ТҮРЕ	LEARNING GOALS
Citytown Meltdown	Digital Game	Explore how changing the temperature of matter affects its state
What is it?	Hands-On Activity	Students identify and describe solids and liquids in their surroundings
Is it Solid? Is it Liquid?	Digital Notebook	Practice distinguishing solids and liquids
The Lake Mistake	Digital Episode	Investigate how matter changes states along with the Sparks Crew
Melt It!	Digital Notebook	Use cause and effect reasoning to explain the process of melting
Changing Materials	Hands-On Activity	Explore how objects react when heated
Solid or Liquid Books!	Digital E-books	Learn the fundamentals of matter and its states

Study Design and Methodology

Four after-school programs in the San Francisco Bay Area participated in the pilot study. Two after-school programs served as the treatment group and implemented the playlists for eight weeks. The other two after-school programs served as the comparison group and were paired with the treatment after-school programs based on their location and grade structure (mixed-grade class or single-grade class).

Treatment and Comparison Conditions

In order to support implementation of the playlists at the treatment after-school programs, TPT led a day-long playlist content training¹ and WestEd led a one-hour research training. During the playlist content training, TPT facilitators provided an overview of the Hero Elementary universe and a brief introduction on the playlist. In addition to introducing the hands-on activities, digital notebooks, and digital games, the facilitators also highlighted the importance of communicating the principles of equity. By the end of the playlist content training, after-school educators were asked to complete the Educator Action Plan for how and when they were planning to complete each playlist in their classroom. The WestEd research training focused on research study expectations, tasks, and logistics.

⁻

¹ Due to educator turnover at Site A, WestEd modified the original day-long content training and a one-hour research training to a four-hour training for a new educator from Site A. For additional details on this WestEd-led educator training, please see Appendix A.

First grade, second grade, and mixed-grade after-school classrooms were encouraged to complete at least four playlists, whereas kindergarten after-school classrooms were encouraged to complete at least three playlists. The eight-week treatment suggested that educators implement each playlist for one to two weeks, three times per week, one hour each time.

Comparison group classrooms behaved according to business-as-usual. Both comparison after-school programs had a schedule that included homework, snacks, outdoor play, and indoor activities. Comparison Site A held indoor activity time in the middle of their afternoon and offered different stations featuring STEM and art activities. Comparison Site B ended their afternoon with indoor activity time or outdoor play throughout the week. Comparison Site B educators planned their indoor activity with the intention of supplementing the students' school curriculum. Comparison Site A did not allow, nor did it desire, technology because the program lacked proper internet security and parents preferred limited technology use. In contrast, Comparison Site B expressed interest in having tablets or Chromebooks to offer their students for educational use but did not currently have the resources necessary to acquire and support the technology.

Instruments

Outcome Measures

ScienceQuest. In accordance with the learning goals of the playlists used in the intervention, measures of student science learning needed to: (a) assess content that would be addressed by the four implemented playlists; (b) integrate core ideas of the Next Generation Science Standards; and (c) be administered with fidelity to both intervention and comparison groups within a time window that would allow sites sufficient time to implement the playlists before the end of the school year. Given the target ages of students receiving the intervention, few pre-existing assessments could achieve this goal. In partnership with 3C, WestEd researchers developed ScienceQuest, a 25-item assessment designed to measure concepts in motion, forces, and properties of matter (see Test Blueprint in the Appendix B).

Each item on ScienceQuest was scored correct or incorrect, with a possible range in total score on the assessment from 0 to 25. Following the pre-implementation administration of ScienceQuest, WestEd researchers investigated the reliability of the assessment, which was found to be acceptable (Cronbach alpha = 0.75). ScienceQuest was administered to students using tablets (the same technology as the playlist activities) while also providing accessibility supports to children.

Emerging STEM Learning Activation Survey. In addition to studying content learning outcomes, the study also investigated the the affective dimension of the intervention – changes in student attitudes towards science and scientific behaviors. This survey is designed by the Learning Activation Lab (http://activationlab.org), and has been used to assess the degree to which a child demonstrates emerging STEM learning activation. The Likert-type scale in the survey consists of a series of facial pictures that represent sad, neither sad nor happy, and happy faces. The survey included a total of 14 items with a total score that could range from 14 to 42. The reliability of the pre-survey was 0.78.

Implementation Measures

Site Observations. Researchers conducted site visits at treatment sites to observe how educators implemented the playlist activities, and at comparison sites to establish a comparative educational context. The observations focused on three key categories of observation: educator behavior, technology performance, and student engagement. Educator behavior captured instruction, lesson structure, and assistance provision. Technology performance included glitches, malfunctions, and other challenges related to technology integration. Well-performing technology was also noted. Student engagement captured students' on/off task behavior and affect.

The Baker Rodrigo Ocumpaugh Monitoring Protocol (BROMP) was used to measure student engagement. BROMP was originally designed for observation of students using educational technology and has been adapted for broader use. When using BROMP, WestEd researchers constructed an ordered list of the students to be observed individually and then recorded a behavior and affect for each student. Researchers observed students sequentially for the duration of the activity, cycling through the ordered list of students as many times as each session allowed. Observations were recorded using a mobile device-compatible application designed specifically for BROMP observation. Observations captured behavior in four categories (on task, on task conversation, off task, and undefined), and affect in five categories (concentrating, bored, confused, frustrated, and undefined).

Educator logs. Educators completed a brief, online survey at the end of each day that they had used the playlists with their class. The log asked about educators' progress in implementing the playlist activities, how long activities took to complete, and details about any problems that arose during implementation. After educators completed a playlist, they had the option to note which playlist activities were successful or challenging, as well as providing general feedback about the playlist through the educator log.

Student focus group/performance tasks. At treatment sites, groups of 5-7 students in each grade level participated in researcher-led focus groups. Students answered questions about new

knowledge they gained from the *Hero Elementary* program, occasions where they used science during the program and at home, and their favorite activity from the program. The focus group typically lasted around 15 minutes.

At both treatment and comparison sites, students from each grade level participated in two short performance tasks in pairs. The first task asked students to make observations about small stones that researchers provided, to sort the rocks into two groups, and to write descriptions of the similarities and differences between the two groups.

For the second task, researchers laid out a tape measure on the ground and had students make predictions about how far a toy car would go when given large and small pushes or pulls. Students were then asked to demonstrate large and small pushes and pulls and to record how far the car went. Lastly, students were asked to draw a picture depicting how they would make the car stop.

Parent focus groups. Parents at each treatment site were invited to participate in researcher-led focus groups that lasted approximately 10 minutes. Parents were asked about their experience in receiving emails or text messages from the *Hero Elementary* program. Focus group questions also gauged parents' level of knowledge about *Hero Elementary* and their child's enjoyment of the program.

Educator interview. Each educator at a treatment site participated in a 30- to 45-minute interview. Interview questions covered educator background and experience level, and information gleaned from watching students interact with the playlist activities. Educators were also asked about ways that they had modified implementation to better fit the needs of their students and to provide opinions and feedback about various components of the program.

Directors of the comparison after-school programs also participated in interviews of a similar length where they were asked about their program's typical schedule, implementation of STEM-related activities, and access to technology.

Study Sample Description

After-school educators were invited to participate in the study through the distribution of informational letters and consent forms through the after-school programs. Parent informational letters and opt-out forms were distributed by after-school educators. Both parent informational letters and opt-out forms were available in English and Spanish. A total of 173 kindergarten to 2nd grade students were recruited to participate in the pilot study, resulting in 86 in the treatment group and 87 in the comparison group. Of the original sample, 168 students across treatment and comparison groups were administered both the pre-assessment and the post-assessment. Students

who did not complete at least 50% of the assessment on either the pre- or post- assessment were removed from the analytic sample, resulting in 161 students in the final analytic sample. The treatment and comparison groups did not differ significantly on their attitudes toward science at baseline as measured by the Emerging STEM Learning Activation Survey. However, the treatment group had higher pre-test scores on ScienceQuest (Table 6).

Table 6. Key Measures at Baseline, by Treatment Condition

Measure	Treatment	Comparison	Difference	<i>p</i> -value	SMD ^a
Emerging Activation Survey					
Mean	35.55	36.51	-0.96	0.184	-0.22
Standard Deviation	4.73	4.34			
N	83	78			
ScienceQuest Pre-Test					
Mean	18.64	16.62	2.02	0.001**	0.48
Standard Deviation	3.42	4.17			
N	83	78			

^{**} Significantly different from zero at the .01 level, two-tailed test. *** Significantly different from zero at the .001 level.

a. SMD refers to Standardized Mean Difference which was calculated by dividing treatment and comparison difference by the comparison group standard deviation of the pre-measure variable.

Data Analysis Methods

WestEd employed a multivariate regression model to determine the effect of playlists on student science knowledge as measured by the ScienceQuest test, and attitudes towards science as measured by the Emerging STEM Learning Activation Survey, after accounting for pre-existing differences in baseline measures. This method is preferred for analyzing pre/post design data because it can eliminate systematic bias and reduce error variance (Bonate, 2000; Dimitrov & Rumrill, 2003). The method also considers regression toward the mean (Bland & Altman, 1994), which refers to the concept that when the first measurement of a variable is an extreme value it will tend to be closer to the average on a later measurement. Given that regression toward the mean is a relatively common phenomenon, the researchers applied a general linear model with pre-tests as covariates in this study to increase statistical power and obtain a more precise and less biased estimate of the group effects (Bonate, 2000; Dimitrov & Rumrill, 2003; Keppel & Wickens, 2004).

As described in the equation below, the model includes variables to account for covariate factors that might impact student performance regardless of receipt of intervention (such as their prior test scores or their grade). The model also includes a variable for site pairings.

 $Outcome_{ij} = \beta_1 Intervention_i + \beta_2 PreScore_i + \sum \beta_G Grade_i + \sum \beta_S SitePair_j + Constant$

Playlist Completion

All educators in the treatment group were asked to complete brief educator logs at the end of every session in which they implemented the playlist. The educator logs attempted to gauge each classroom's progress in completing each playlist, which digital and/or hands-on activities were implemented during each session, and any technical difficulties or obstacles encountered during that session. Based on educator log data, the kindergarten classroom completed three playlists, while the first grade, second grade, and mixed-grade classrooms completed all four playlists. No sites completed the "optional" playlist on *Heating and Cooling* (see Table 7).

Table 7. Playlist Completion by Classroom

Classroom	# of Participating Students in Each Classroom	Playlists Completed
Kindergarten	17	Changing Motion Make it Fast, Make it Slow (Kindergarten)
		Pushes and Pulls
1st Grade	20	Changing Motion
		Classifying Matter
		Pushes and Pulls
		Solids and Liquids ^a
2nd Grade	22	Changing Motion
		Classifying Matter
		Pushes and Pulls
		Solids and Liquids
Mixed-Grade	K: 12	Changing Motion
Classroom	1st Grade: 8	Classifying Matter
	2nd Grade: 13	Pushes and Pulls
		Solids and Liquids

a. Educator log data were not received for the implementation of this playlist at this site's classroom, but other sources of data confirm that this playlist was completed.

The educator logs also asked educators to indicate which digital and hands-on activities the students participated in during that session, and the approximate total minutes that the students participated in either the digital or the hands-on activities during the session. The educators provided their responses using the following time ranges: 10 minutes or less; 11-20 minutes; 21-30 minutes; 31-40 minutes; or 41-50 minutes. During analyses, the time ranges were converted into an approximate number of total minutes by taking the mid-point of each of the time ranges. According to the educator log records received, Treatment Site A spent a total of 1385 minutes on digital activities and a total of 325 minutes on hands-on activities, while Treatment Site B spent a total of 395 minutes on digital activities and 280 minutes on hands-on activities.

Student Behavior and Affect

Researchers coded student behavior and affect using the BROMP protocol during four activity sessions. Table 8 summarizes the observed sessions.

Table 8. Observed playlist sessions—BROMP

	SITE	GRADE	PLAYLIST	ACTIVITIES
Session 1	Site A	Mixed (K-2)	Solid or Liquid	Video (individual viewing) and digital game
Session 2	Site B	2 nd	Classifying Matter	Video (co-viewing)
Session 3	Site A	Mixed (K-2)	Solid or Liquid	E-books and digital notebook
Session 4	Site B	Mixed (K-2)	Solid or Liquid	Digital notebook

Session 1

Of 155 total observed instances of students watching a video (individual viewing) and using the digital game, 85% of coded behavior was on task. Similarly, observed students exhibited 129 instances of concentration out of 155 coded affective states (83%), indicating that students were engaged and focused throughout the session. There was only a single coded moment of boredom (<1%), two coded moments of confusion (1%), and two coded moments of frustration (1%). While students might have experience confusion or frustration while transitioning between activities, low rates of those affective states indicated that students understood how to navigate between playlist activities on the tablet. In addition, 21 observations (14%) received behavior codes indicating that a student was away from the activity. These resulted primarily from early parental pick-ups.

Session 2

Researchers coded a total of 102 behavior and affective states during Session 2, in which students watched a *Hero Elementary* video (co-viewing). Students were, for the most part, on task and concentrating (94%). There was some conversation during the video, though students mostly remained focused even while speaking (14% of on-task behavior included conversation). Overall, students appeared to understand and to be engaged with the video, as only one instance each of boredom and confusion were recorded.

Session 3

Like in Sessions 1 and 2, students were mostly on task and concentrating (87% of 111 total observations) while using the e-books and digital notebook. While 12% of on-task behavior included conversation, students mostly worked individually. There were six recorded instances of confusion (6%), mostly related to the notebook prompt. Only one instance of boredom was observed (<1%).

Session 4

In Session 4, students were far more frequently off task, with only 52% of recorded observations (128 total) showing on-task behavior. This is likely because students were using the digital notebook for the second day in a row and students wanted to do other activities. Educators attempted to keep students focused on the notebook task, but many students switched to other playlist activities—especially the digital game. These students displayed an affect of concentration but were coded as off task (31%) because educators intended them to be using the notebook. An additional 15% exhibited a bored affect.

Classroom Integration

Digital Game

Educators varied in their approaches to introducing the digital game to students. The majority of the educators did not demonstrate how to play the digital game. They allowed students to begin exploring the levels on their own and provided aid to students who raised their hands. One kindergarten educator was able to provide ample scaffolding to her kindergarten class the first time that they used the digital game. The educator held the tablet upright on her lap and explained how to navigate the digital environment. Kindergarteners sat on the carpet and watched as the educator walked them through the first few levels of the game. The educator set up the tables around the room ahead of time and dismissed the students one at a time to distribute the tablets. After dismissing all the students from the carpet, she circulated – providing assistance as students logged in and started playing.

Researchers observed students playing the game individually, in pairs, and in groups of up to four students. No debriefings of the digital games were observed. Educators allowed students to play the digital game up to the end of the activity session in observed classes.

E-Books

During observed sessions, students across classrooms were required to read the e-books individually, with their headphones on. Educators provided instructions about how to access the e-books and what to do when students finished reading. However, they did not contextualize the content of the e-books for students. They did not discuss the science concepts that the e-books covered or how they related to the playlist core ideas.

One exception to the above occurred during a class in which many of the students were struggling to connect their tablets to the internet. The educator decided to use the e-books as a whole-class activity. She read from the tablet to the class, holding the screen up to show pictures. She reported that this went well and allowed for discussion of new vocabulary that the e-books introduced, such as inertia and friction.

Co-Viewing Hero Elementary Videos

In all three observed classrooms that attempted to co-view the video, logistics proved to be a barrier. All three educators were unfamiliar with how to use the projectors and speakers at their site and required extensive on-site support to set up the video and projector. One of the educators chose not to use the projector and instead had students gather around a tablet to watch.

While the logistics of co-viewing were challenging, one educator found that the co-viewing was helpful. She found that watching the video as a class allowed her to pause the episode and incorporate content from the co-viewing guide. "Because we were watching it together and we were able to stop it, to make predictions or anything like that, so they can understand what was going on. So, they weren't just watching it, they were actually watching it and trying to pick out the problem or the solution."

Not every educator used the co-viewing guide. In one observed debriefing, an educator asked students to raise their hand and to share what they learned from the video. Many students were eager to share. The educator gave each student a turn to speak but did little to respond to their comments or reiterate key ideas. Researchers also observed an activity session in which students watched the video and then transitioned to a different playlist activity without any discussion of its content.

Science Notebook

Researchers observed some students using the notebook to explore science and to respond to the notebook prompts. However, many students did not engage with the science prompts at all and used the notebooks in a way that was unrelated to playlist science content. Some educators reported that students were allowed "free draw" time after they had completed the assigned activity in the notebook.

Researchers observed that educators often introduced the notebook activity by focusing on the logistics rather than science content. Educators did not appear to deliberately connect the notebook to other playlist activities as a tool for recording data, observations, or reflections. One educator did report in the interview that she used the notebook as a tool for reflection after each Hands-on activity (HOA). She commented, "Whenever we'd [do] the hands on, they right away knew that we were gonna do a notebook. So, then we would talk about the notebook together [with the HOA]." While that educator reported discussing the HOAs as part of the notebook activity, no debriefings occurred during observed sessions.

Most educator support pertained to technical glitches or other malfunctions. Although researchers did not observe educators providing instructions about how to use the notebook's

features, many students explored the tool and figured out how to take pictures, record their voice, and navigate the illustrator.

Hands-On Activities

Educators reported that they modified the length or difficulty of some HOAs to fit the needs of their students. One educator reported ending the sorting activity before the recommended time because students finished quickly. Another educator modified the sorting activity to make it more challenging for the same reason by having students find additional objects in the classroom to sort into the specified categories. The other educator described modifying *One More Push* because it was too easy for older students. To make it more challenging, she modified the activity and asked older students to measure and record the distance traveled for each object that they pushed. Some other educators changed the mini-golf activity, because it felt repetitive to students. Instead of waiting until the third time they did the activity to allow students to build their own courses, educators had students create obstacles in the second session. They then rotated groups around the room so that students could navigate each other's courses.

While these modifications demonstrate planning and thoughtfulness in the way educators implemented the HOAs, researchers also observed a session in which a lack of planning was evident. The educator asked a researcher who was on site for an observation where the instructions for *Changing Materials* were. She glanced through the instructions that the researcher identified but did not prepare any materials before implementing the activity with students.

Connections

During observed sessions, introductions to the playlist activities focused on instructions and logistics, in-game assistance focused on logistics and troubleshooting, and activity debriefings were often unstructured reviews of the activity. In all three types of educator support, facilitating connections was not a primary focus and was not consistently included in educators' scaffolding.

One educator recalled a student connecting the e-book to the digital game; another remembered a student connecting the science content in the *Classifying Matter* playlist to the science in *Solids and Liquids*. However, most educators reported during interviews that they did not observe their students making connections between playlists, between activities, or to their own lives. This was borne out in observations of playlist activity implementation, as researchers did not observe students making connections.

Classroom Management

Researchers observed that two of the four treatment classes encountered issues with classroom management. In these classes, students were frequently off task and educators struggled to redirect students to the playlist activities. Students frequently left their seats and moved around

the room, played with materials that were unrelated to the playlist activities (e.g., sports equipment, gaming cards), and engaged with the playlist activities in ways that were inconsistent with the learning goals.

In general, educators lacked classroom management strategies. When multiple educators were in a classroom and facilitated activities, they were able to manage off-task behaviors. However, when there was only one educator facilitating the activity, off-task behavior occasionally impeded implementation. Off-task behaviors were commonly linked to other implementation challenges. For example, researchers observed that educators had greater difficulty managing student behaviors when the class was experiencing technology issues. Students were likely to distract each other if they finished an activity quickly. These distracting behaviors were particularly common among older students who were not challenged during HOAs.

Age Appropriateness of Content

Educators' perceptions of the playlists' age appropriateness varied across sites. Educators at one site felt that the content was difficult for kindergarten students, but appropriate for first- and second-grade students. The kindergarten educator commented, "For the most part, the structure....is good. It's just they're just too young and little that [the activities] might not be as interesting to them.... But I do see they're able to be sufficient for first grader/second graders." The first-grade educator at this site noticed that students were focused when they interacted with the digital activities, and never appeared distracted. This suggested that the activities were at the appropriate level for first-grade students. The second-grade educator at this site found that some of the HOAs took longer than the allotted time, indicating that they were sufficiently complex and challenging for second-grade students.

However, the educators at the other site believed that the content was easy for all of their students, especially the second-grade students. Activities generally took much less time than what was allotted, indicating students might not have been challenged. One educator at this site commented, "I noticed that even our first graders and kindergartners, it was easy for them. I was kind of surprised that they.... finished that in five or ten minutes, when it should have been, like, twenty minutes." Educators at this site remarked that the content would be better suited to preschoolers or students with IEPs. They felt that this was especially true for the playlist, *Pushes and Pulls*. One educator at this site reported, "This playlist was too easy and not much of a challenge."

Digital Games

At one site, educators felt that the digital games challenged their students appropriately. The first-grade educator at this site noted how much students enjoyed the digital games, and how students were able to work together to understand the games. The educator commented, "Yeah, they did, they really enjoyed the game, as any kid would. Yeah, they would, like, try to, like, if

one kid didn't get it the other kids would try to show them how to do it. So, which I thought, that was nice." One student mentioned that the heating and cooling game was difficult, but fun. This suggests that the games were at the appropriate level for these students and provided them with opportunities to learn from one another. The kindergarten educator at this site appreciated that the *Pushes and Pulls* game gave her students the opportunity to think and problem solving independently. She commented, "I think them actually figuring out what to do in this scenario where they had to push or pull, I think that *[sic]* kind of what made them think on their own. 'Okay then, I need to push something, pull this' and then it's, like, each day they would come into this situation where they had to pull or push things and kind of know what to do and what to expect."

At the other site, educators found that the digital games were too easy across all grade levels. Though some students struggled with the *Pushes and Pulls* digital game, an educator at this site commented, "Most of them found it very easy or too easy." The educator recommended adding additional levels to the game "because a lot of the kids finished the game early."

Hands-On Activities

Educators at one site tended to feel that the HOAs were too difficult for the kindergarten students. During the *Changing Motion* playlist, for example, the kindergarten educator felt that the overall structure wasn't suitable for kindergarteners, and commented, "The hands-on game became a little challenging at times and I strongly believe it's because they are still five years old. Taking turns and being patient is still something they are getting used to."

However, the HOAs tended to go quickly at the other site across grade levels. Educators at this site had to continuously make modifications to extend the activities in order to challenge their students. One educator commented, "The hands-on activity *One More Push* was very simple and too easy for [the] majority of children. Therefore, I added some extra things to do during the project to keep the children engaged, specifically the older children." This educator recommended extending the activities and creating different levels for each grade level.

At both sites, educators agreed that the most successful activities were the ones that provided students with opportunities to explore independently. Educators at both sites mentioned that *Push, Slide, Bump* was a successful learning experience. The format of this activity allowed for greater variation, making it adaptable across grade levels. The second-grade classroom at one site spent more time doing this activity because students were so absorbed in creating obstacles. An educator at the other site mentioned that students at this site also spent additional time on this activity: "They liked it because they were able to modify the route that the object must travel to get to the finish line. I noticed that when kids are given an instruction on how to do it they will get bored pretty fast. They liked when they were able to input their own ideas into the activity."

E-Book

The e-books may have been too difficult for the kindergarten students to read independently at one site. Due to technical issues (detailed in a later section) the kindergarten educator had students read the e-book as a group in the *Pushes and Pulls* playlist. She felt that this instructional strategy worked out better than having students read independently "because all of the vocabulary was new to them, such as inertia and friction."

Educators at the other site thought that the e-books were at the appropriate level for their kindergarten and first-grade students. However, they felt that the e-books were too easy for the second-grade students:

The second grader [sic], they are pretty swift and quick on catching on things. ... when I went back even to talk to them, I said, 'Wait a minute, you guys got done already with the book, did you even read it?' And they're, like, 'Yeah, we did.' And I was, like, 'Okay, tell me what the book is about.' And then they were able to tell me what they read on the e-book.

Technical Issues

Technical issues created a significant barrier to implementation and student engagement. The biggest barriers were: (1) the log-in process; (2) the logistics of using devices such as the mi-fi and the projector; and (3) various glitches in digital activities. These technical issues frustrated students to the point of disengagement with playlist activities.

Login Process

Across sites, educators struggled with the scanning process required for login. The login process created barriers to implementation, and sometimes kept educators from being able to complete activities in the time that they had planned. These complications caused some educators to forgo the scanning process altogether, preferring to manually enter the passwords instead:

Of course, the login was always a challenge at first. One of the simple things, like, for example, the cue cards, the placement of where the whole punch on the string went or right over the barcode. But we had a hard time with that. And, of course, with the lighting issue, and, of course, sometimes with technology...some of the codes weren't working, like when we try to manually enter it.

The only challenge for me was when I wanted them to do something and we couldn't get something done. For example, like, let's get the book going and we cannot go back and read the book or we have, like, a log-in issues [sic] so we couldn't start the activity because we couldn't get you logged in and stuff like that.

I usually just wouldn't even try scanning it, I would just input it manually....I think [the scanning process] just took too long. You had to have, like, do [a] double take to make sure it went through or maybe it wouldn't go through at all."

Logistics

Internet. Across sites, educators reported that tablets did not reliably connect to the internet and loading time was often significant. One educator reported, "Some of the tablets weren't connecting to the internet or connection, even with refreshing the page, exiting the page entirely." Another educator mentioned, "The internet wasn't working on all the iPads, so in order to save time I ended up doing the reading as a group and we discuss[sic] every page together after it was read aloud."

These instances may have been issues with either the Playlist site or with the internet itself and was likely a mixture of the two. Undeniably, these problems impacted the way that educators were able to implement the playlists.

Projector. The logistics of setting up devices for video co-viewing posed a barrier for educators. Educators needed support to understand the different cords, mi-fi devices, and speaker set up. A kindergarten educator opted to have the students crowd around one tablet instead of using the projector. A first-grade educator commented, "The projector, that was super hard for me 'cause I didn't really get where to connect all the cables." A second-grade educator was able to get the projector to work, but only with significant on-site assistance from researchers during every viewing. Set-up took significant time, effort, and coordination on the educators' part and proved to be a challenge. This impacted the way that the playlists were implemented; educators were not always able to show the video in the time they had planned or ended up showing videos out of the prescribed order of playlists.

Glitches

Aside from trouble using the scanning feature of the log-in/QR codes, educators experienced a number of glitches during digital activities that impeded implementation, student engagement, and completion of the playlist. Students sometimes were so frustrated with the app itself that it kept them from engaging with the activities. These issues can be found below, broken down by activity.

Games. Educators reported frequent glitches with the *Pushes and Pulls* digital game. At both sites, educators reported days on which none of the students were able to play through the entirety of the game.

About 90% of the games were freezing today. Many kids finished most of their activities and when they clicked to play the game it just kept giving the loading signal. We restarted the tablets and still the game would not load.

Science Notebook. Although educators did not report glitches as frequently during the use of the notebook as they did during the game, educators reported several issues with the notebook. Educators repeatedly reported that the notebook would become "extremely large" upon opening it in the app. Educators were unable to resize the image and had to log out and restart the program to fix the problem. Educators also reported that the screen sometimes went white, and that students were unable to erase their work.

For awhile, the notebook for the tablets wasn't working so they had to keep reading and do other ones they could and two of the students now no longer want to do the tablets at all and they just choose to do their homework.

E-Book. Students experienced persistent difficulties in accessing the e-books. Some educators reported that students were unable to access a second book and thought that it was a technical issue. Some other educators reported that students were sometimes unable to access any e-books.

There were some kids that weren't able to go on the e-book sometimes. Like, they would click on it and it would go to something weird, like some random page. Then, I would tell them to go to another computer and sometimes it'll work, but sometimes it'll do the same thing. So, yeah, they wouldn't be able to read the book.

A researcher saw one such issue during an observation: when a student tried to access the e-book, the screen began flashing seemingly random pictures. When the educator explained that the student would need to switch devices, the student groaned, "I don't want to." After the student finally got on another computer and began to engage with the e-book, he realized that everyone else had already finished. He then clicked through the e-book without reading it. This demonstrates the frustration and impatience that students had with the technical issues, and the ways that these problems prevented students from engaging with the playlist activities.

Co-Viewing Videos. Educators at both sites reported that videos did not always function properly. Educators reported that, "The video would get blurry even though it already loaded. Some *[sic]* were clear but a lot of the other kids' were blurry." One video was shared on two playlists. When students were watching the video for the second time, researchers observed that some students appeared bored and were lying down during the video.

Icons. Students seemed motivated to complete activities in order to watch the lightning bolts turn green. However, the tracking was not functioning properly, and students were frustrated and wanted to stop using the playlist.

Many kids finished the e-book, drawing, and hands-on activity, but their icons didn't change to green, indicating that they completed. This made many kids frustrated. Some kids were saying 'That's not fair, I finished but my icon didn't change to green.' Another few kids finished their drawings but their icon didn't turn green and when they clicked on it again it made them redo the activity as if they didn't do it. That made the kids feels [sic] frustrated and didn't [sic] want to continue.

Impact Results

The results of the multivariate linear regression indicate that the intervention may have been associated with a minor gain in children's science knowledge (as measured by the ScienceQuest assessment); however, the effect cannot be considered statistically significant for an α of 0.05 (p \sim 0.38). Adjusted mean differences on the post-test measure of ScienceQuest showed that students who received the intervention exceeded their comparison peers in total score (point estimate of 0.41) which corresponded to an effect size of 0.10² but which had a confidence interval of [-0.51, 1.33]. The results did not show a significant difference in post-implementation science attitudes (as measured by Emerging STEM Learning Activation Survey), with an α of 0.05 (point estimate of 0.39; p \sim 0.55; see Table 9).

Table 9. Student Science and STEM Attitude Outcomes

	Adjuste	d Mean					
Impact measure	Treatment (Standard Error)	Control (Standard Error)	Difference	p- value	95% confidence interval	sample size	effect size ^a
STEM Survey							
	35.59 (0.45)	35.20 (0.47)	0.39	0.551	-0.90-1.68	161	0.08
ScienceQuest							
	19.00 (0.32)	18.59 (0.34)	0.41	0.382	-0.51-1.33	161	0.10

a. Effect size was calculated by dividing treatment and comparison difference by the standard deviation of the outcome measure among the comparison group.

 $^{^{2}}$ Effect size was calculated by dividing treatment and comparison difference by the standard deviation of the outcome measure among the comparison group(SD = 4.38).

Although quantitative analysis of overall assessment and survey results did not show significant differences between treatment and comparison groups, analysis of individual items on ScienceQuest showed some promising results. These results are further unpacked in Appendix C. In particular, WestEd examined differential improvement on *Science Quest* items based on receipt of treatment. Table 10, below, highlights items which saw statistically significant difference in post-intervention improvement. Five items, corresponding to the playlists "Changing Motion" and "Classifying Matter", saw intervention-receiving students improve beyond their comparative peers.

Table 10. Item-Level Performance Change by Intervention Condition

Science Quest Item			Improvement on Post-test		Pearson's chi-squared	
Item Name	Corresponding Playlist	Condition	0^{a}	1 ^b	statistic	p-value
H00X03	Changing Motion					
		control	7	70	3.359	0.067†
		treatment	2	81		
H00X07	Changing Motion					
		control	7	70	3.359	0.067†
		treatment	2	81		
H00X09	Changing Motion					
		control	22	53	3.332	0.068†
		treatment	14	68		
H00X21	Classifying Matter					
		control	26	48	2.713	0.100†
		treatment	19	63		
H00X23	Classifying Matter					
		control	36	39	7.537	0.006**
		treatment	22	60	1 . 1 . 1	

a. 0 indicates students who regressed from their pretest performance or provided an incorrect answer during both tests b. 1 indicates students who improved upon their pretest performance or provided the correct answer during both tests Note: \dagger , *, or ** respectively denote significant difference at a p-value of .1, .05, and .01 levels.

Additionally, our observations and interviews with educators and students showed some potential impact of the playlists on student science attitudes. When observing the *Classifying Matter* playlist that used the flashlight, researchers found that the use of the flashlight was very exciting to students. At first, the flashlight seemed to be a distraction. As time went on, however, it was clear that students were able to use the flashlight as a tool to make discoveries. Students were interested in the way that light moved (or did not move) through different objects. It did not seem that they were making many discoveries about the objects themselves; rather, they were making discoveries about light and were eager to continue this. The fact that

students had ownership over their discoveries seemed to resonate with two boys who did not want to stop the activity. (Interestingly, these same two boys were fairly disengaged with the digital activities.) In addition, educators reported that some students were engaged in the science activities and would love to explore more science.

One of the kids knew we were doing the science theme that day, and then...he told me if I could call his grandma to pick him up later because he wanted to do the theme. Then, I called her and she's, like, "Oh, yeah, he really likes that science theme."

I think, for example for my kinder, they didn't really do much science for the after-school program...so, they were always bringing stuff to art, so that's all we ever did, art, art, art. So then, when they did this it was, like "Okay, science is something that I do like." Or that is interesting to me. So, then they started asking me, like, "Okay," they actually were the ones to bring it up to me. "Oh, okay. Let's do science." So, I knew that they were interested in it when they kept asking me about it almost every day.

We prepared them, we asked them, "So, who here loves science?"....We try to turn it around and tell them, "Well, you guys will be doing, you'll be learning about how you classify matter or how to push and pull blocks...or you'll be watching video about learning about science." I think the way the kids that didn't want to learn about science wanted to learn after we said that learning science is fun. It can be fun in the class or electronically. When they heard electronically they were like, "Oh, are we getting to use tablets?" They got super excited because they could learn science but still have a balance between technology and deduction.

Well, what I think it exposes ...the children to, is that there are other ways to learn science...besides just the classroom... especially in this era or generation, I think it's really important for children to understand technology use, but then also understand how it can be beneficial in the educational sense.

I think with this program, most of our kids like to try it out. They want to see, they want to feel, they get the feel of it and if they are feeling it, then of course, the interest for it goes upwards. I feel like right now after we did this program, I think maybe one or two kids are going to....We've got a light bulb towards more exposure to science. For the other kids, I think we didn't do the light bulb but not because of the interest, just because maybe they didn't get more time with it.

Conclusion

The results indicated that educators were able to integrate the playlists into their after-school programs. Educators were able to complete three to four playlists in eight weeks. They were able

to implement each playlist for one to two weeks, three times per week, and one hour each time. Students enjoyed playing the digital games, which directly reinforced science concepts covered throughout the playlists. Teachers reported some potential impact of the playlists on their students' science attitudes. The findings from the study also showed students' knowledge of science was maintained at the same level, as measured by ScienceQuest.

Many of the participating educators did not have adequate classroom management trainings and lacked experience in implementing youth development curricula. We expect that the technical glitches will be solved through further configuration and development of the digital assets. In order to implement the playlists in informal learning environments, it is also important to provide more systematic trainings to educators with the focus on understanding the playlist contents, best practices for implementing each type of activity in informal learning environments, and technology trouble shooting. In addition, supporting materials on how to modify the contents to better meet the needs of different age groups is important for the success of playlist implementation.

Reference

- Basu SJ, Barton AC. (2007). Developing a sustained interest in science among urban minority youth. *Journal of Research in Science Teaching*. 44:466–489. doi: 10.1002/tea.20143.
- Bland, J.M. & Altman, D.G. (1994). Statistic Notes: Regression towards the mean. *British Medical Journal*, 308, (6942): 1499.
- Bonate, P.L. Analysis of Pretest-Posttest Designs. Chapman & Hall, 2000.
- Dimitrov, D.M. & Rumrill, P.D. *Pretest-Posttest Designs and Measurement of Change*. Work 20, 159-165, IOS Press, 2003.
- Dabney KP, Tai RH, Almarode JT, Miller-Friedmann JL, Sonnert G, Sadler PM, Hazari Z. (2012). Out-of-school time science activities and their association with career interest in STEM. *International Journal of Science Education, Part B.*;2(1):63–79.
- Dimitrov, D. M., & Rumrill, P. D. (2003). Pretest-posttest designs and measurement of change. *Work (Reading, Mass.)*, 20(2), 159–165.
- Keppel, G. & Wickens, T.D. (2004). *Design and analysis: A researcher's handbook* (4th ed). Upper Saddle River, NJ: Pearson Prentice Hall.
- Ozel M, Caglak S, Erdogan M. (2013). Are affective factors a good predictor of science achievement? Examining the role of affective factors based on PISA 2006. *Learning and Individual Differences*. 24:73–82. doi: 10.1016/j.lindif.2012.09.006.
- Peterson K., Fix S. (2007). Afterschool Advantage: Powerful New Learning Opportunities. Retrieved from http://www.statewideafterschoolnetworks.net/afterschool-advantage-powerful-new-learning-opportunities.
- Pierce KM., & Bolt DM., & Vandell DL. (2010). Specific Features of After-School Program Quality: Associations with Children's Functioning in Middle Childhood. *American journal of community psychology.* 45, 381-93. 10.1007/s10464-010-9304-2.
- Vandell DL, Reisner ER, Pierce KM. (2007). Outcomes Linked to High-Quality Afterschool Programs: Longitudinal Findings from the Study of Promising Afterschool Programs. Policy Studies Associates, Inc.

Appendix A: WestEd Educator Training

WestEd provided a pre-study training to one educator at Site A that joined the study as a replacement lead for the kindergarten classroom. The presentation was based on the training that TPT provided but tailored to the educator's circumstances as a replacement teacher and modified to emphasize certain teaching practices. It contained five parts, summarized below. See included slides for further details.

Introduction

The introduction presented the *Hero Elementary* transmedia universe, situating the playlists as one component of many interrelated science learning tools. Presenters introduced each member of the Sparks Crew and explained the education goals of the playlists and the broader *Hero Elementary* initiative.

Superpowers of Science

Presenters defined the term "Superpowers of Science" and discussed how the Superpowers fit into the playlists and the *Hero Elementary* video episodes. Slides identify and explain seven key Superpowers.

Hero Elementary Playlists

In this section, presenters defined the term "playlist." They walked the audience through the different types of activities that make up a *Hero Elementary* playlist.

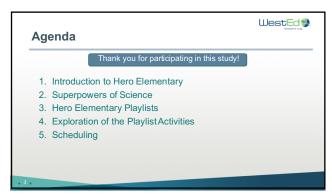
Exploration of the Playlist Activities

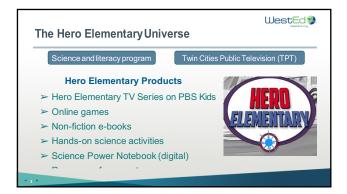
With guidance and support from WestEd staff, the educator set up and played through an HOA and explored the digital activities. For each activity, WestEd staff discussed how to introduce the activity to students in a way that connected to the playlist's core idea, incorporate language related to the Superpowers of Science into the activity, and debrief the activity with students.

Scheduling

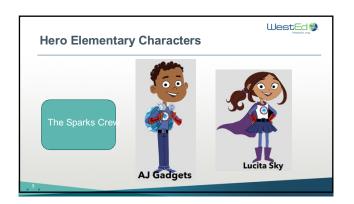
Presenters discussed playlist implementation logistics. This included how long to spend on each activity, how to sequence activities and playlists, and repetition of activities within a playlist.

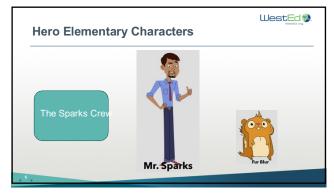


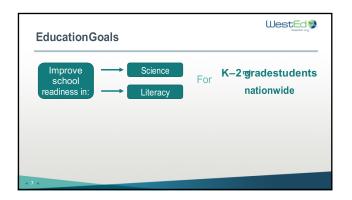


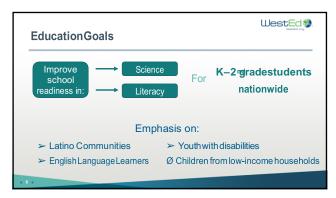








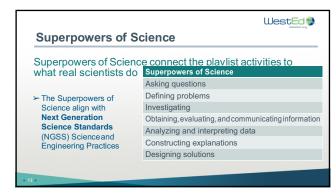


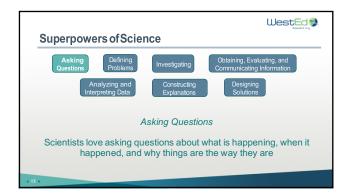


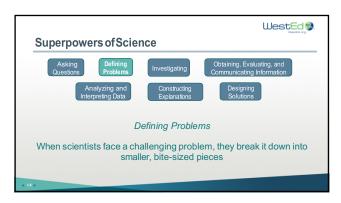


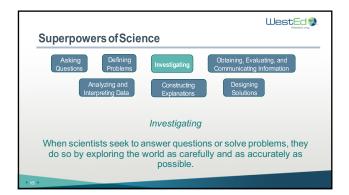


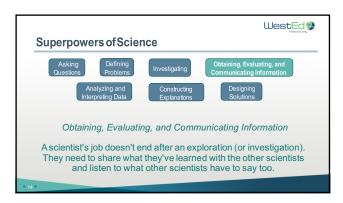


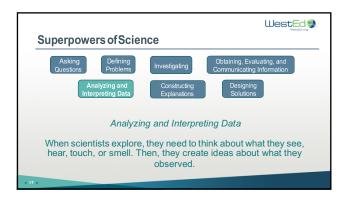


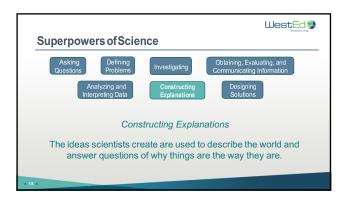


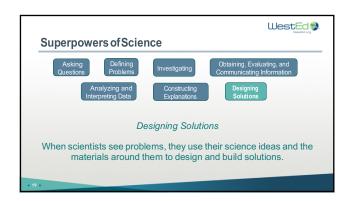


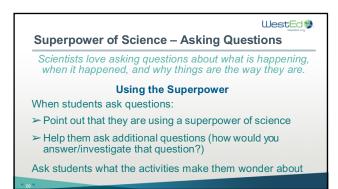




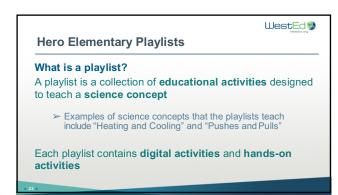








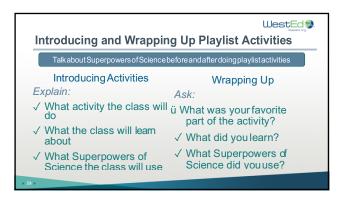








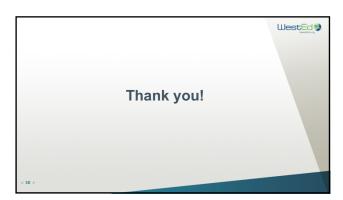












Appendix B: ScienceQuest Test Blueprint

Item Name	Item Grade Level	Primary Target Playlist	Pushes and Pulls	Make it fast, make it slow	Changing Motion	Classifying Matter	Solid and Liquid	DCI: PS1.A	DCI: PS1.B	DCI: PS2.A	DCI: PS2.B	DCI: PS3.C
H00X27	K	Make it fast, make it slow	Х	х	Х					Х		Х
H00X17	K	Make it fast, make it slow		х	Х						Х	х
H00X17	K	Changing Motion		X	X						X	X
H00X06	K	Make it fast, make it slow		X	X						X	X
H00X12	K	Changing Motion		X	X					x	Α	X
H00X01	K	Pushes & Pulls	х	Α	Λ					X		^
H00X01	K	Pushes & Pulls	X							X		
H00X11	K	Pushes & Pulls	X	х	Х					X	Х	х
H00X07	K	Changing Motion	^	Α	X					^	X	X
H00X07	K	Changing Motion			X						X	X
H00X05	K	Make it fast, make it slow	х	х	X					х	Λ	X
H00X05	K	Make it fast, make it slow	X	X	X					X	x	X
H00X16	K	Changing Motion	X	X	X					X	Α	X
	K	Pushes & Pulls	X	Λ	Λ					X	X	^
H00X04 H00X19		Solid and Liquid	^				X		x		Λ	
H00X19	1	Pushes & Pulls	Х		Х		Α		^	x		х
H00X13 H00X29	K		Α		Α		v	х	х			^
H00X29	2	Solid and Liquid Solid and Liquid					X	X				
H00X25	1	•				x	Α		x			
H00X25	2	Classifying Matter				X		X	X			
	2	Classifying Matter						X				
H00X21	2	Classifying Matter				Х	**	X	X			
H00X14	1	Solid and Liquid					Х	Х				
H00X23	1	Classifying Matter				Х			X			
H00X22	2	Classifying Matter				X		Х	Х			
H00X20	1	Solid and Liquid				Х	X	Х	Х			

- DCI: PS1.A "Matter exists as different substances that have observable different properties.

 Different properties are suited to different purposes. Objects can be built up from smaller parts."
- DCI: PS1.B "Heating and cooling substances cause changes that are sometimes reversible and sometimes not."
- DCI: PS2.A "Pushes and pulls can have different strengths and directions. Pushing or pulling on an object can change the speed or direction of its motion and can start or stop it."
- DCI: PS2.B "When objects touch or collide, they push on one another and can change motion."
- DCI: PS3.C "Bigger pushes and pulls cause bigger changes in an object's motion or shape"

Appendix C: Item-level Impact Results

Item-level Performance Change

Science Test Questions										
Question	Group		No .1	Reg	ression ²	Improv	vement ³	Per	rfect ⁴	
		impro N	vement ¹ %	N	%	N	%	N	%	Total (N)
H00X01										
	Comparison	7	9%	11	14%	7	9%	52	68%	77
	Intervention	1	1%	11	13%	10	12%	61	73%	83
	all	8	5%	22	14%	17	11%	113	71%	160
H00X02										
	Comparison	2	3%	4	5%	2	3%	69	90%	77
	Intervention	0	0%	2	2%	1	1%	80	96%	83
	all	2	1%	6	4%	3	2%	149	93%	160
H00X03										
	Comparison	1	1%	6	8%	7	9%	63	82%	77
	Intervention	0	0%	2	2%	2	2%	79	95%	83
	all	1	1%	8	5%	9	6%	142	89%	160
H00X04										
	Comparison	2	3%	8	10%	9	12%	58	75%	77
	Intervention	2	2%	7	9%	9	11%	63	78%	81
	all	4	3%	15	9%	18	11%	121	77%	158
H00X05									0.107	
	Comparison	2	3%	1	1%	4	5%	70	91%	77
	Intervention	0	0%	2	2%	1	1%	80	96%	83
H00X06	all	2	1%	3	2%	5	3%	150	94%	160
H00X06	Comparison	25	32%	13	17%	14	18%	25	32%	77
	Intervention	25	25%	16	19%	20	24%	26	31%	83
	all	46	29%	29	18%	34	21%	51	32%	160
H00X07	411	10	= 2 / 0		10,0		21/0		J2 / 0	100
22002107	Comparison	1	1%	6	8%	6	8%	64	83%	77
	Intervention	0	0%	2	2%	1	1%	80	96%	83

¹ No improvement indicates that student answered the question incorrectly on both pre- and post-tests

² Regression indicates that student answered the question correctly on pre-test but incorrectly on post-test ³ Improvement indicates that student answered the question incorrectly on pre-test and correctly on post-test

⁴ Perfect indicates that student answered the question correctly on both pre- and post-tests

	all	1	1%	8	5%	7	4%	144	90%	160
H00X09										
	Comparison	14	19%	8	11%	11	15%	42	56%	75
	Intervention	7	9%	7	9%	14	17%	54	66%	82
	all	21	13%	15	10%	25	16%	96	61%	157
H00X10										
	Comparison	17	22%	4	5%	10	13%	47	61%	77
	Intervention	11	13%	13	16%	8	10%	51	61%	83
	all	28	18%	17	11%	18	11%	98	61%	160
H00X11										
	Comparison	4	5%	0	0%	8	10%	64	83%	77
	Intervention	2	2%	1	1%	3	4%	77	93%	83
	all	6	4%	1	1%	11	7%	141	88%	160
H00X12										
	Comparison	8	10%	12	16%	11	14%	46	60%	77
	Intervention	7	8%	12	14%	11	13%	53	64%	83
	all	15	9%	24	15%	22	14%	99	62%	160
H00X13										
	Comparison	2	3%	8	10%	5	6%	62	81%	77
	Intervention	1	1%	5	6%	10	12%	66	80%	83
	all	3	2%	13	8%	15	9%	128	80%	160
H00X14										
	Comparison	24	31%	9	12%	11	14%	31	40%	77
	Intervention	18	22%	9	11%	16	19%	39	47%	83
	all	42	26%	18	11%	27	17%	70	44%	160
H00X15										
	Comparison	7	9%	2	3%	17	22%	48	62%	77
	Intervention	4	5%	4	5%	10	12%	65	78%	83
	all	11	7%	6	4%	27	17%	113	71%	160
H00X17										
	Comparison	4	5%	5	6%	14	18%	55	71%	77
	Intervention	1	1%	7	8%	2	2%	73	88%	83
****	all	5	3%	12	8%	16	10%	128	80%	160
H00X19										
	Comparison	5	6%	8	10%	13	17%	51	66%	77
	Intervention	4	5%	4	5%	11	13%	64	77%	83
	all	9	6%	12	8%	24	15%	115	72%	160
H00X20	G		000/		100/	4.0	A F O (
	Comparison	22	29%	14	18%	19	25%	21	27%	77
	Intervention	22	27%	15	18%	14	17%	28	34%	83

	all	44	28%	29	18%	33	21%	49	31%	160
H00X21										
	Comparison	16	21%	10	13%	18	23%	30	39%	77
	Intervention	13	16%	6	7%	17	20%	46	55%	83
	all	29	18%	16	10%	35	22%	76	48%	160
H00X22										
	Comparison	26	34%	7	9%	17	22%	25	32%	77
	Intervention	20	24%	7	8%	13	16%	41	49%	83
	all	46	29%	14	9%	30	19%	66	41%	160
H00X23										
	Comparison	23	30%	13	17%	10	13%	29	38%	77
	Intervention	16	19%	6	7%	8	10%	52	63%	83
	all	39	24%	19	12%	18	11%	81	51%	160
H00X24										
	Comparison	27	35%	14	18%	19	25%	17	22%	77
	Intervention	30	36%	11	13%	18	22%	24	29%	83
	all	57	36%	25	16%	37	23%	41	26%	160
H00X25										
	Comparison	19	25%	12	16%	22	29%	22	29%	77
	Intervention	22	27%	10	12%	21	25%	30	36%	83
	all	41	26%	22	14%	43	27%	52	33%	160
H00X26										
	Comparison	28	36%	8	10%	21	27%	21	27%	77
	Intervention	19	23%	11	13%	21	25%	32	39%	83
	all	47	29%	19	12%	42	26%	53	33%	160
H00X27										
	Comparison	10	13%	7	9%	17	22%	44	57%	77
	Intervention	7	8%	8	10%	13	16%	55	66%	83
	all	17	11%	15	9%	30	19%	99	62%	160
H00X29										
	Comparison	20	26%	7	9%	18	23%	29	38%	77
	Intervention	17	20%	13	16%	16	19%	36	43%	83
	all	37	23%	20	13%	34	21%	65	41%	160

Comparison Group Cross-Tabulation

Paired Cross-tal	bulation	S	McNemar's chi-squared			
Post.test						
Question	Pre.test	0	1	statistic	p-value	

H00X01					
	0	7	7	0.889	0.346
	1	11	52		
H00X02					
	0	2	2	0.667	0.414
	1	4	69		
H00X03					
	0	1	7	0.077	0.782
	1	6	63		
H00X04					
	0	2	9	0.059	0.808
	1	8	58		
H00X05				1.000	0.100
	0	2	4	1.800	0.180
11001106	1	1	70		
H00X06	0	25	14	0.027	0.947
	1	13	25	0.037	0.847
H00X07	1	13	23		
1100/X07	0	1	6	0.000	1.000
	1	6	64	0.000	1.000
H00X09					
	0	14	11	0.474	0.491
	1	8	42		
H00X10					
	0	17	10	2.571	0.109
	1	4	47		
H00X11					
	0	4	8	8.000	0.005**
	1	0	64		
H00X12		0	1.1	0.042	0.025
	0	8	11	0.043	0.835
1100V12	1	12	46		
H00X13	0	2	5	0.692	0.405
	1	8	62	0.072	0.703
H00X14	1		02		
1100211	0	24	11	0.200	0.655
	1	9	31		
	-				

H00X15					
	0	7	17	11.842	0.001**
	1	2	48		
H00X17					
	0	4	14	4.263	0.039*
	1	5	55		
H00X19					
	0	5	13	1.190	0.275
*****	1	8	51		
H00X20	0	22	10	0.250	0.611
	0	22	19	0.258	0.611
H00X21	1	14	21		
Π00λ21	0	16	18	2.286	0.131
	1	10	30	2.200	0.131
H00X22	1	10	30		
11007322	0	26	17	4.167	0.041*
	1	7	25		
H00X23					
	0	23	10	0.391	0.532
	1	13	29		
H00X24					
	0	27	19	0.758	0.384
	1	14	17		
H00X25	0	4.0	2.2		0.006
	0	19	22	2.941	0.086
11003/27	1	12	22		
H00X26	0	10	21	5 929	0.016*
	1	28 8	21 21	5.828	0.016*
H00X27	1	0	21		
1100/12/	0	10	17	4.167	0.041*
	1	7	44		VVV 11
H00X29					
***************************************	0	20	18	4.840	0.028*
	1	7	29		

Intervention Group Cross-Tabulation

Post Presentation Presentation	THE VEHICION ON	oup Cross-rabi	racic	711		
Hound	Paired Cross-tabulation	ons			McNemar's chi-	-squared
H00X01			Pos	t.test		
HOUXO2	Question	Pre.test	0	1	statistic	p-value
HOUXO2	H00X01					
H00X02 0 0 1 0.333 0.564 H00X03 0 0 2 0.000 1.000 1 2 79 H00X04 0 2 9 0.250 0.617 1 7 63 H00X05 0 0 1 0.333 0.564 H00X06 0 0 1 0.333 0.564 H00X06 0 21 20 0.444 0.505 1 16 26 H00X07 0 0 1 0.333 0.564 H00X07 0 0 1 0.333 0.564 H00X09 1 1 2 80 H00X09 1 1 7 54 H00X10 0 11 8 1.190 0.275 H00X11 0 2 3 1.000 0.317 H00X11 H00X12 0 7 11 77 H00X12		0	1	10	0.048	0.827
HOOX03		1	11	61		
HOOXO3 HOOXO4 HOOXO4 O	H00X02					
H00X03 0 0 2 0.000 1.000 H00X04 0 2 9 0.250 0.617 1 7 63 H00X05 0 0 1 0.333 0.564 1 2 80 H00X06 0 21 20 0.444 0.505 1 16 26 H00X07 0 0 1 0.333 0.564 H00X09 0 7 14 2.333 0.564 H00X10 0 11 8 1.190 0.275 1 13 51 H00X11 0 2 3 1.000 0.317 H00X12 0 7 11 77 H00X12					0.333	0.564
H00X04 H00X05 H00X06 D D D D D D D D D D D D D D D D D D		1	2	80		
H00X04 H00X05 0 2 9 0.250 0.617 1 7 63 H00X05 0 0 1 0.333 0.564 1 2 80 H00X06 0 21 20 0.444 0.505 1 16 26 H00X07 0 0 1 0.333 0.564 1 2 80 H00X09 0 7 14 2.333 0.564 H00X10 0 11 8 1.190 0.275 1 13 51 H00X11 0 2 3 1.000 0.317 H00X11 H00X12 0 7 11 1 77 H00X12	H00X03					
H00X04 0 2 9 0.250 0.617 H00X05 0 0 1 0.333 0.564 H00X06 0 21 20 0.444 0.505 1 16 26 H00X07 0 0 1 0.333 0.564 H00X09 0 7 14 2.333 0.564 H00X10 0 11 8 1.190 0.275 1 13 51 H00X11 0 2 3 1.000 0.317 H00X11 0 7 11 77 H00X12 0 7 11 77 H00X12		0			0.000	1.000
H00X05 H00X06 H00X06 0 0 1 0.333 0.564 1 2 80 H00X06 0 21 20 0.444 0.505 1 16 26 H00X07 0 0 1 0.333 0.564 H00X09 0 7 14 2.333 0.127 H00X10 0 11 8 1.190 0.275 1 13 51 H00X11 0 2 3 1.000 0.317 H00X12 H00X12 0 7 11 77 H00X12		1	2	79		
H00X05 0 0 1 0.333 0.564 H00X06 0 21 20 0.444 0.505 1 16 26 H00X07 0 0 1 0.333 0.564 H00X09 0 0 1 0.333 0.564 1 2 80 H00X09 0 7 14 2.333 0.127 H00X10 0 11 8 1.190 0.275 1 13 51 H00X11 0 2 3 1.000 0.317 H00X12 0 7 11 77 H00X12	H00X04		2	0	0.250	0.617
H00X05 0 0 1 0.333 0.564 H00X06 0 21 20 0.444 0.505 1 16 26 H00X07 0 0 1 0.333 0.564 1 2 80 H00X09 0 7 14 2.333 0.127 1 7 54 H00X10 0 11 8 1.190 0.275 1 13 51 H00X11 0 2 3 1.000 0.317 H00X12 0 7 11 77 H00X12					0.250	0.61/
H00X06 H00X07 0 21 20 0.444 0.505 1 16 26 H00X07 0 0 1 0.333 0.564 H00X09 0 0 1 0.333 0.564 H00X09 0 7 14 2.333 0.127 1 7 54 H00X10 0 11 8 1.190 0.275 1 13 51 H00X11 0 2 3 1.000 0.317 1 1 77 H00X12 0 7 11 0.043 0.835	H00V05	1	/	03		
H00X06 H00X07 0 21 20 0.444 0.505 1 16 26 H00X07 0 0 1 0.333 0.564 1 2 80 H00X09 0 7 14 2.333 0.127 1 7 54 H00X10 0 11 8 1.190 0.275 1 13 51 H00X11 0 2 3 1.000 0.317 1 1 77 H00X12 0 7 11 0.043 0.835	1100203	0	0	1	0.333	0.564
H00X07 H00X07 0 0 1 0.333 0.564 1 2 80 H00X09 0 7 14 2.333 0.127 1 7 54 H00X10 0 11 8 1.190 0.275 1 13 51 H00X11 0 2 3 1.000 0.317 1 1 77 H00X12 0 7 11 0.043 0.835						
H00X07 0 0 1 0.333 0.564 H00X09 0 7 14 2.333 0.127 1 7 54 H00X10 0 11 8 1.190 0.275 1 13 51 H00X11 0 2 3 1.000 0.317 1 1 77 H00X12 0 7 11 0.043 0.835	H00X06					
H00X07 0 0 1 0.333 0.564 H00X09 0 7 14 2.333 0.127 1 7 54 H00X10 0 11 8 1.190 0.275 1 13 51 H00X11 0 2 3 1.000 0.317 1 1 77 H00X12 0 7 11 0.043 0.835 1 12 53		0	21	20	0.444	0.505
H00X09 0 7 14 2.333 0.127 1 7 54 H00X10 0 11 8 1.190 0.275 1 13 51 H00X11 0 2 3 1.000 0.317 1 1 77 H00X12 0 7 11 0.043 0.835 1 12 53		1	16	26		
H00X09 0 7 14 2.333 0.127 1 7 54 H00X10 0 11 8 1.190 0.275 1 13 51 H00X11 0 2 3 1.000 0.317 1 1 77 H00X12 0 7 11 0.043 0.835 1 12 53	H00X07				0.000	0.564
H00X09 0 7 14 2.333 0.127 1 7 54 H00X10 0 11 8 1.190 0.275 1 13 51 H00X11 0 2 3 1.000 0.317 1 1 77 H00X12 0 7 11 0.043 0.835 1 12 53					0.333	0.564
H00X10 0 7 14 2.333 0.127 H00X10 0 11 8 1.190 0.275 1 13 51 H00X11 0 2 3 1.000 0.317 1 1 77 H00X12 0 7 11 0.043 0.835 1 12 53	H00V00	1		80		
H00X10 0 11 8 1.190 0.275 1 13 51 H00X11 0 2 3 1.000 0.317 1 1 77 H00X12 0 7 11 0.043 0.835 1 12 53	1100209	0	7	14	2.333	0.127
H00X11 0 11 8 1.190 0.275 1 13 51 H00X11 0 2 3 1.000 0.317 1 1 77 H00X12 0 7 11 0.043 0.835 1 12 53						
H00X11 0 11 8 1.190 0.275 1 13 51 H00X11 0 2 3 1.000 0.317 1 1 77 H00X12 0 7 11 0.043 0.835 1 12 53	H00X10					
H00X11 0 2 3 1.000 0.317 1 1 77 H00X12 0 7 11 0.043 0.835 1 12 53		0	11		1.190	0.275
H00X12 0 2 3 1.000 0.317 1 1 77 H00X12 0 7 11 0.043 0.835 1 12 53		1	13	51		
H00X12 0 7 11 0.043 0.835 1 12 53	H00X11			2	1.000	0.215
H00X12 0 7 11 0.043 0.835 1 12 53					1.000	0.317
0 7 11 0.043 0.835 1 12 53	H00V12	1	1	/ /		
1 12 53	П00Х12	0	7	11	0.043	0.835
					0.015	0.000
	H00X13					

	0	1	10	1.667	0.197
	1	5	66		
H00X14					
	0	18	16	1.960	0.162
	1	9	39		
H00X15					
1100/X13	0	4	10	2.571	0.109
	1	4	65	2.3 / 1	0.107
11003/17	1	7	0.5		
H00X17	0	1	2	2.770	0.007
	0	1	2	2.778	0.096
	1	7	73		
H00X19					
	0	4	11	3.267	0.071
	1	4	64		
H00X20					
	0	22	14	0.034	0.853
	1	15	28		
H00X21	_				
1100/121	0	13	17	5.261	0.022*
	1	6	46	3.201	0.022
~~~~~	1	U	40		
H00X22		•	1.0	4.000	0.400
H00X22	0	20	13	1.800	0.180
H00X22	0	20 7	13 41	1.800	0.180
H00X22 H00X23				1.800	0.180
				1.800 0.286	0.180
	1	7	41		
H00X23	0	7	8		
	0 1	7 16 6	8	0.286	0.593
H00X23	0 1	7 16 6	8 52 18		
H00X23	0 1	7 16 6	8 52	0.286	0.593
H00X23	0 1	7 16 6 30 11	8 52 18 24	0.286	0.593
H00X23	0 1 0 0	7 16 6 30 11	8 52 18 24	0.286	0.593
H00X23 H00X24 H00X25	0 1	7 16 6 30 11	8 52 18 24	0.286	0.593
H00X23	0 1 0 1	7 16 6 30 11 22 10	8 52 18 24 21 30	0.286 1.690 3.903	0.593 0.194 0.048
H00X23 H00X24 H00X25	0 1 0 1 0 1	7 16 6 30 11 22 10	8 52 18 24 21 30	0.286	0.593
H00X23 H00X24 H00X25	0 1 0 1	7 16 6 30 11 22 10	8 52 18 24 21 30	0.286 1.690 3.903	0.593 0.194 0.048
H00X23 H00X24 H00X25	0 1 0 1 0 1	7 16 6 30 11 22 10	8 52 18 24 21 30	0.286 1.690 3.903	0.593 0.194 0.048
H00X23 H00X24 H00X25 H00X26	0 1 0 1 0 1	7 16 6 30 11 22 10	8 52 18 24 21 30	0.286 1.690 3.903	0.593 0.194 0.048
H00X23 H00X24 H00X25 H00X26	0 1 0 1 0 1 0 1	7 16 6 30 11 22 10	41 8 52 18 24 21 30 21 32	0.286 1.690 3.903	0.593 0.194 0.048
H00X23 H00X24 H00X25 H00X26 H00X27	0 1 0 1 0 1 0 1	7 16 6 30 11 22 10 19 11	41 8 52 18 24 21 30 21 32	0.286 1.690 3.903	0.593 0.194 0.048
H00X23 H00X24 H00X25 H00X26	0 1 0 1 0 1 0 1	7 16 6 30 11 22 10 19 11	41 8 52 18 24 21 30 21 32	0.286 1.690 3.903	0.593 0.194 0.048

1 | 13 | 36 |

# All Conditions (Comparison and Treatment) Cross-Tabulation

<b>Paired Cross-</b>	-tabulations			McNemai	r's chi-squared
		Po	st.test		
Question	Pre.test	0	1	statistic	p-value
H00X01					
	0	8	17	0.641	0.423
	1	22	113		
H00X02			2	1.000	0.215
	0	2	3	1.000	0.317
H00X03	1	6	149		
П00Л03	0	1	9	0.059	0.808
	1	8	142	0.039	0.000
H00X04	1	O	172		
11002101	0	4	17	0.273	0.602
	1	22	113	1.2,5	0.002
H00X05					
	0	2	5	0.500	0.480
	1	3	150		
H00X06					
	0	46	34	0.397	0.529
	1	29	51		
H00X07					
	0	1	7	0.067	0.796
11003/00	1	8	144		
H00X09	0	21	25	2.500	0.114
	0	21	25	2.500	0.114
H00X10	1	15	96		
1100/110	0	28	18	0.029	0.866
	1	17	98	0.02)	0.000
H00X11		. ,			
	0	6	11	8.333	0.004**
	1	1	141		
H00X12					
	0	15	22	0.087	0.768
	1	24	99		
H00X13					
	0	3	15	0.143	0.705
11001/14	1	13	128		
H00X14	0	40	27	1.000	0.100
	0	42	27	1.800	0.180
	1	18	70		

H00X15					
	0	11	27	13.364	0.000***
	1	6	113		
H00X17					
	0	5	16	0.571	0.450
	1	12	128		
H00X19					
	0	9	24	4.000	0.046*
	1	12	115		
H00X20					
	0	44	33	0.258	0.611
	1	29	49		
H00X21					
	0	29	35	7.078	0.008*
	1	16	76		
H00X22					
	0	46	30	5.818	0.016*
	1	14	66		
H00X23					
	0	39	18	0.027	0.869
	1	19	81		
H00X24					
	0	57	37	2.323	0.128
	1	25	41		
H00X25					
	0	41	43	6.785	0.009*
	1	22	52		
H00X26					
	0	47	42	8.672	0.003**
	1	19	53		
H00X27					
	0	17	30	5.000	0.025*
	1	15	99		
H00X29					
	0	17	30	3.630	0.057
	1	15	99		

# Post-Test Performance Cross-Tabulation by Condition

	errormance (				
Cross-tabul	lations			Pearson's ch	ii-squared
		Improvemen	t on Post Test		
Question	Condition	$0^{\mathrm{a}}$	1 ^b	statistic	p-value
H00X01					
	control	18	59	2.086	0.149
	treatment	12	71		
H00X02					
	control	6	71	2.436	0.119
	treatment	2	81		
H00X03					
	control	7	70	3.359	0.067†
	treatment	2	81		
H00X04					
	control	10	67	0.131	0.717
	treatment	9	72		
H00X05					
	control	3	74	0.292	0.589
	treatment	2	81		
H00X06		2.0	2.0	0.265	0.716
	control	38	39	0.365	0.546
*******	treatment	37	46		
H00X07	oomtwol	7	70	3.359	0.047÷
	control treatment	2	81	3.359	0.067†
H00X09	ti eatment	<u> </u>	01		
Hooko	control	22	53	3.332	0.068†
	treatment	14	68		00000
H00X10					
	control	21	57	0.079	0.778
	treatment	24	59		
H00X11					
	control	4	72	0.256	0.613
	treatment	3	80		
H00X12					
	control	20	57	0.206	0.650
	treatment	19	64		
H00X13					

	, 1	10	67	1 411	0.225
	control	10	67	1.411	0.235
	treatment	6	76		
H00X14					
	control	33	42	2.034	0.154
	treatment	27	55		
H00X15					
	control	9	65	0.258	0.611
	treatment	8	75		
H00X17					
	control	9	69	0.154	0.695
	treatment	8	75		
H00X19					
	control	13	64	1.839	0.175
	treatment	8	75		
H00X20					
	control	36	40	0.004	0.947
	treatment	37	42		
H00X21					
11001121	control	26	48	2.713	0.100†
	treatment	19	63		33233
H00X22					
11007122	control	33	42	1.872	0.171
	treatment	27	54	1.072	0.171
H00X23	troutmont	27	31		
11002125	control	36	39	7.537	0.006**
	treatment	22	60	7.007	0.000
H00X24	treatment		00		
1100X24	control	41	36	0.237	0.626
	treatment	41	42	0.237	0.020
1100725	treatment	41	42		
H00X25	aantual	31	44	0.127	0.722
	control			0.127	0.722
HOOMAG	treatment	32	51		
H00X26	1	26	40	1 ((5	0.107
	control	36	42	1.665	0.197
*******	treatment	30	53		
H00X27		1.77	61	0.250	0.554
	control	17	61	0.350	0.554
	treatment	15	68		
H00X29					
	control	27	47	0.000	0.990

treatment	30	52	
-----------	----	----	--

a. 0 indicates students who regressed from their pretest performance or provided an incorrect answer during both tests b. 1 indicates students who improved upon their pretest performance or provided the correct answer during both tests Note: †, *, or ** respectively denote significant difference at a p-value of .1, .05, and .01 levels.

# Performance Task Analysis

The performance tasks were conducted with a total of 35 students across the treatment and control group: 11 Kindergarteners, 9 first-graders, 15 second-graders (Table 1). The performance task comprised questions evaluating the students' grasp of the topics, "classifying matter" (task 1) and "pushes and pulls" (task 2). Evaluations were performed by different researchers from WestEd. For the purposes of our analysis, we removed kindergarten students from the task 1 analyses, as kindergarten students were not given the "classifying matter" playlist.

Table 1a. N by grade level and condition - Task 1

	Treatment Group	Control Group	Total
Kindergarten	0	0	0
ı st Grade	6	3	9
2 nd Grade	10	5	15
Total	16	8	24

Table 1b. N by grade level and condition - Task 2

	Treatment Group	Control Group	Total
Kindergarten	7	4	11
ı st Grade	6	3	9
2 nd Grade	10	5	15
Total	23	12	35

We conducted the analyses in 4 different scenarios. In the original scenario, we retained the original scoring in which task 1 was scored out of 16 and task 2 was scored out of 12.

In Scenario 1, all sub-scores of "4" were recoded to "3". Hence, the maximum possible score for task 1 in scenario 1 is 12 and the maximum possible score for task 2 in scenario 1 is 9.

In Scenario 2, all sub-scores of "4", "3", and "2" were recoded to 1, and sub-scores of "1" and "0" were recoded to "0". Hence, the maximum possible score for task 1 in scenario 2 is 4, and the maximum possible score for task 2 in scenario 2 is 3.

In Scenario 3, all sub-scores of "4", and "3" were recoded to 2, and sub-scores of "2" were recoded to "1", and sub-scores of "0" and "1" were recoded to "0". Hence, the maximum possible score for task 1 in scenario 3 is 8, and the maximum possible score for task 2 in scenario 3 is 6.

We found that the treatment, for the sample size tested, did not significantly increase the performance of students on this task in any of the scenario analyses.

We present the results of our analyses below.

Table 2. Average total score by condition, task, and scenarios

	Original Scenario		Scenario 1		Scenario 2		Scenario 3	
	Average Score - Task 1 (Total Out of 14)	Average Score Task 2 (Total Out of 12)	Average Score - Task 1 (Total Out of 12	Average Score Task 2 (Total Out of 9)	Average Score - Task 1 (Total Out of 4)	Average Score Task 2 (Total Out of 3)	Average Score - Task 1 (Total Out of 8)	Average Score Task 2 (Total Out of 6)
Tx	12.7	10.0	10.56	8.22	3.63	2.91	6.56	5.26
Cx	13.3	11.2	10.88	8.67	3.63	2.92	6.88	5.75